

BAŞKALARI TARAFINDAN OLUŞTURULAN İSTATİSTİKSEL SONUÇ VEYA YORUMLARI TARTIŞMA

İstatistiksel veriler karar verme süreçlerinde hayati bir rol oynasa da, bu verilerin sunduğu sonuçlar her zaman mutlak gerçeği yansıtmayabilir. Bir analizin doğruluğu, çevresel, ekonomik veya bireysel değişkenlerin (mevsimsellik, gelir düzeyi, psikolojik durum vb.) sürece dahil edilip edilmediğine bağlıdır. Bu nedenle, önümüze sunulan verileri körü körüne kabul etmek yerine, hangi bağlamda hazırlandıklarını sorgulayan **eleştirel ve objektif bir süzgeçten geçirmek** gerekir.

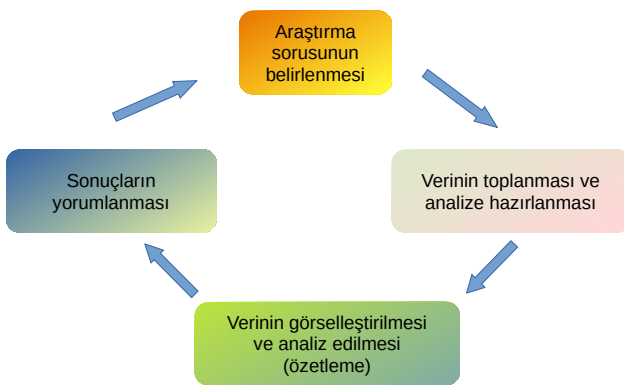
Temel Noktalar:

Bağlamın Önemi: Veriler, toplandıkları koşullardan bağımsız değerlendirilemez.

Yanıtıcı Sonuçlar: Önemli değişkenlerin dışarıda bırakılması, istatistikleri manipülatif veya eksik hale getirir.

Sorgulayıcı Yaklaşım: Güvenilir kararlar alabilmek için başkalarının sunduğu yorumlara karşı şüpheci ve tarafsız kalmak esastır.

İSTATİSTİKSEL ARAŞTIRMA SÜRECİ AŞAMALARI



Örnek...1 :

A futbol takımı yöneticileri, 2015 yılından başlayarak 2020 sonuna kadar takımlarının maç sayıları ve ortalama seyirci sayılarını (bin kişi) internet sitelerinde şekildeki tabloda sunmuşlardır.

Yıl	2014	2015	2016	2017	2018	2019	2020
Maç sayısı	40	44	42	46	42	46	40
Ortalama seyirci sayısı	20	24	22	26	24	22	15

Yukarıdaki verileri inceleyen bir öğrenci bu verileri kullanarak her yıl maçların ortalama ziyaretçi sayılarının azaldığına dair bir yorumda bulunmuştur.

Öğrencinin yorumu hatalıysa bu hatanın istatistiksel araştırma sürecinin hangi aşamasında olabileceğini belirtiniz.

Veri Analizi: Tablo incelendiğinde 2014'ten 2015'e (20 den 24 e) ve 2016'dan 2017'ye (22 den 26 ya) bir artış olduğu görülmektedir. Öğrenci, verinin tamamına bakmak yerine sadece son yılları veya genel eğilimi yanlış yorumlamış eya verileri dikkatli incelememiş..Cevap: Hata, Sonuçların yorumlanması aşamasındadır.

İstatistiksel Araştırma Süreci ve Hata Aşaması

İstatistiksel bir araştırma süreci temel olarak şu adımlardan oluşur:

Araştırma sorusu oluşturma.

Uygun verileri toplama.

Verileri tablo veya grafiğe dönüştürerek düzenleme.

Verileri analiz etme ve yorumlama.

Bu problemde veriler doğru bir şekilde toplanmış ve tabloya (3. aşama) hatasız aktarılmış. Öğrencinin hatası, önündeki hazır veriyi yanlış değerlendirip hatalı bir sonuç çıkarılmasından kaynaklanıyor.

Bu yüzden hata; istatistiksel araştırma sürecinin "Verileri analiz etme ve yorumlama" aşamasındadır.

Örnek...2 :

2024 yılında sosyal medya platformlarına reklam vermeyi düşünen bir kişi bu platformların son 5 yıldaki kullanıcı sayılarına (tabloda milyon kişi) bakarak popüleritelerindeki değişimi ve kullanıcı sayılarının fazlalığına göre kararını vermek istiyor. Bu amaçla topladığı verilerle aşağıdaki tabloyu oluşturuyor ve 2024 yılı için tahmini kullanıcı sayısını son sütüne yazıyor.

Yıl	2019	2020	2021	2022	2023	2024
A	1,8	2,4	2,6	2,4	2,2	2,8
B	1,2	2	2,2	2	2,2	2,1
C	0,5	1	1,5	1,6	1,8	2
D	0,2	0,4	0,7	0,8	0,9	0,6

Tablodaki verilere göre, bu kişi hangi platformun/platformların 2024 yılı için kullanıcı sayısını tahminini hatalı/yanlış hesaplamış olabilir ?

Platform A: 2.2'den 2.8'e sıçrama yapmış. Trend son 3 yıldır düşüştükten (2.6 dan 2.4 e 2.2 ye) bu artış verilerle uyumsuzdur. Düşüşteki bir platformun aniden rekor kırmasını beklemek istatistiksel olarak oldukça yanlış bir yaklaşım olur.

Platform D: 0.9'dan 0.6'ya düşmüş. Trend sürekli artıştıkten (0.2den \$ 0. , 0.7, 0.8, 0.9) bu düşüş mantıksız görünmektedir. Sürekli yükselen bir grafiğin durduk yere geçmişteki (2021 öncesi) seviyelerine çakılacağını öngörmek mantıklı durmuyor.

Cevap: A ve D platformlarının tahminleri mevcut trendlere aykırıdır.

MERKEZİ EĞİLİM ÖLÇÜLERİ

1. ARİTMETİK ORTALAMA

Bir dizide terimler toplamının toplam veri adetine bölerek buluruz. Aritmetik ortalama, veri grubundaki verilerin denge noktasıdır.

Örnek...3 :

Bir otoparktaki araç sayısı haftanın farklı günleri için 10, 20, 30, 22, 34, 120 ve 115 olarak sayılmıştır. Buna göre, bu hafta için günde ortalama kaç araç bu parkta bulunmaktadır?

$$\{10 + 20 + 30 + 22 + 34 + 120 + 115\} / \{7\} = 351 / 7 \text{ yaklaşık } 50.14$$

2. MEDYAN (ORTANCA)

Dizinin terimleri büyükten küçüğe ya da küçükten büyüğe doğru sıralandığında baştan ve sondan eşit uzaklıktaki sayıya medyan (ortanca) denir.

Eğer veri dizisi tek elemanlıysa ortadaki veri medyandır.

Eğer veri dizisi çift elemanlıysa ortada bulunan iki verinin aritmetik ortalamasıdır.

Örnek...4 :

Veri grupları için medyana hesaplayınız.

- a) 12, 16, 60, 17, 7, 81, 12
b) 52, 26, 6, 7, 8, 12

a) 12, 16, 60, 17, 7, 81, 12

b) 52, 26, 6, 7, 8, 12

Sıralı: 7, 12, 16, 17, 60, 81.

Sıralı: 6, 7, 8, 12, 26, 52.

Medyan: 16

Medyan: $\{8 + 12\} / \{2\} = 10$

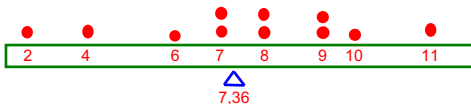
Örnek...5 :

Bir futbol okulunda eğitim gören A takımına ait futbolcular her antrenman sonrası 10 ar adet penaltı çekmektedirler. Gole çevirdikleri penaltı atışları üzerinden performansları değerlendirilmek istenmektedir.

A grubunda eğitim alan futbolcuların 11 sporcunun gole çevirdikleri penaltı atış sayıları 6, 8, 7, 8, 4, 9, 11, 2, 10, 9, 7 olarak belirtilmiştir.

Antrenör, takım oyuncularından aritmetik ortalamasının üzerinde gol atanların performanslarının yüksek olduğunu söylemiştir. A grubu antrenörünün yaklaşımını istatistiksel açıdan değerlendiriniz.

Analiz: Veriler (6, 8, 7, 8, 4, 9, 11, 2, 10, 9, 7) toplandığında 81 eder; ortalama $81 / 11$ yaklaşık 7.36 olur. Değerlendirme: Aritmetik ortalama, grubun genel başarısını temsil eden bir "denge noktasıdır". Ancak 2 ve 4 gibi uç değerler (alt uç değerler) ortalamayı aşağı çeker. Yine de bir futbol takımı için genel seviyeyi belirlemede makul bir yöntemdir.



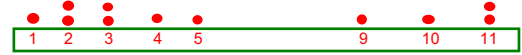
Örnek...6 :

Bir basket okulunda eğitim gören B takımına ait basketçiler her antrenman sonrası 10 ar adet serbest atış (faul atışı) atışı atmaktadırlar. Baskete çevirdikleri serbest atışlar üzerinden performansları değerlendirilmek istenmektedir.

B grubunda eğitim alan 11 basketçinin sayıya çevirdikleri serbest atış sayıları 3, 2, 4, 1, 5, 9, 11, 11, 10, 3, 2 olarak belirtilmiştir.

Antrenör, takım oyuncularından ortanca değerinin üzerinde basketi sayıya çevirenlerin performanslarının yüksek olduğunu söylemiştir. B grubu antrenörünün yaklaşımını istatistiksel açıdan değerlendiriniz.

Sıralı Veri: 1, 2, 2, 3, 3, 4, 5, 9, 10, 11, 11. Verilerin ortalaması yaklaşık 5,54 tür. Medyan: 4. Değerlendirme: Medyan kullanımı, 11 gibi uç değerlerin (üst uç değerlerin) ortalamayı saptırmasını engeller. Takımı tam ortadan ikiye bölerek "standart" bir oyuncunun ne yapması gerektiğini daha sağlıklı gösterir. Ortanca değeri, veri grubunun yoğunlaştığı yeri göstermede aritmetik ortalamaya göre daha uygun bir merkezi eğilim ölçüsü olduğu söylenebilir.



MERKEZİ YAYILIM ÖLÇÜLERİ

Dizideki terimlerin birbirine yakınlığı ya da uzaklığı hakkında bilgi veren ölçülerdir.

1) Üst Uç Değer

Verilerin en büyük olanıdır.

2) Alt Uç Değer

Verilerin en küçük olanıdır

3) Açıklık (Aralık, Ranj)

Veri grubundaki en büyük ve en küçük değer arasındaki farktır

Örnek...7 :

Bir balıkçı sattığı iki farklı balık türü için son 1 haftalık sürede sattığı bu türlere ait karşılaştırma yapmak istemiştir. Yaptığı hesaplamalarda satılan A balığı ve B balığı miktarlarının aritmetik ortalama değerlerinin neredeyse aynı, medyanlarının ise tamamen aynı olduğunu görmüştür. Buna göre, balıkçının daha istikrarlı satış sayısına sahip balığı bulmak için hangi istatistiksel kriteri ölçü alması beklenebilir?

Cevap: Açıklık (Aralık) veya Yayılım Ölçüleri. Aritmetik ortalamalar aynıysa, verilerin birbirine ne kadar yakın olduğunu gösteren açıklık değeri küçük olan balık türü daha istikrarlıdır.

3. MOD (TEPE DEĞER)

Bir dizide en çok tekrarlanan sayıya mod (tepe değeri) denir. Bazı durumlarda ise mod olmayabilir. Bazı durumlarda birden çok olabilir. Veri grubunda her değer farklı ise mod bulunmaz.

Veri grubunda bir terimin tekrar sayısına o terimin frekansı denir.

Örnek...8 :

Giyim eşyası üretiminde en çok satılan numaralar ve bedenlerin dikkate alınırken hangi istatistiksel veri ölçüt alınarak yapılması muhtemeldir?

Mod (Tepe Değer). Çünkü üretimde "en çok tekrarlanan" (en çok tercih edilen) ölçülerin bilinmesi gerekir

Örnek...9 :

Bir internet mağazasında satılan üç farklı akıllı telefon markasının dört yıllık satış miktarlarını gösteren tablo aşağıda verilmiştir.

Yıl	2019	2020	2021	2022
A	1200	1600	1400	1500
B	800	1400	1200	1200
C	600	900	1200	1000

a) Buna göre telefonların satış miktarları hangi istatistiksel araçlarla özetleyebilir?

b) Satış miktarı bakımından diğerlerine göre daha istikrarlı olan marka hangisidir?

Bu veriler hem zaman içindeki değişimi hem de markalar arası farkı gösterdiği için şu araçlar kullanılabilir:Merkezi Eğilim Ölçüleri:Aritmetik Ortalama (\bar{x}): Her markanın yıllık ortalama satış performansını belirlemek için.Medyan (Ortanca): Veri setindeki aşırı artış veya azalışların (uç değerlerin) etkisini minimize ederek orta noktayı görmek için.Yayılım (Değişim) Ölçüleri:Açıklık (Range): En yüksek ve en düşük satış arasındaki farkı bularak genel oynaklığı görmek için.Standart Sapma (σ): Satışların ortalamadan ne kadar uzaklaştığını, yani risk ve istikrar durumunu ölçmek için en sağlıklı araçtır.Görselleştirme Araçları:Çizgi Grafiği: 4 yıllık trendi (yükseliş/düşüş) izlemek için en ideal yöntemdir.Sütun Grafiği: Markaların toplam pazar paylarını karşılaştırmak için kullanılabilir.

A markasının satış rakamları arasındaki farklar gerek açıklık olsun gerekse (varyans) ve standart sapması, diğer iki markaya göre daha düşüktür. B ve C markalarında yıllık bazda daha büyük dalgalanmalar (örneğin B'deki 600 birimlik ani artış gibi) görülmektedir.Sonuç: Satış miktarı bakımından diğerlerine göre en istikrarlı olan marka A markasıdır.A markası "istikrar abidesi" gibi duruyor, B ve C ise biraz daha volatil (oynak) bir grafik çizmiş.

Merkezî eğilim ölçüleri veri grubunun merkezini temsil eden ölçülerdir.

• Merkezî eğilim ölçülerinin her biri, verilerin hangi değerler etrafında toplandığını gösterirken yayılım ölçüleri de verilerin birbirinden ne kadar uzak olduğunu gösteren ölçülerdir.

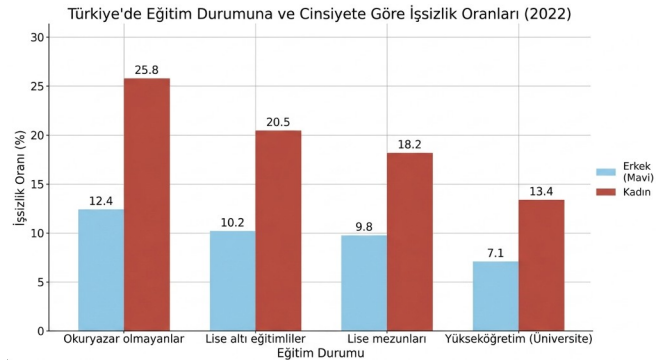
Örnek...10 :

Türkiye'de Eğitim Durumuna Göre İşsizlik Oranları

Eğitim Durumu	Erkek İşsizlik (%)	Kadın İşsizlik (%)
Okuryazar olmayanlar	12,4	25,8
Lise altı eğitilmişler	10,2	20,5
Lise mezunları	9,8	18,2
Yükseköğretim (Üniversite)	7,1	13,4

Yukarıdaki tabloda, bir ülkedeki eğitim durumuna ve cinsiyete göre işsizlik oranları (%) verilmiştir.

Bir ekonomist, bu tabloyu inceledikten sonra alttaki grafiği oluşturmuştur.



Grafiğe bakarak bireylerin iş bulma şansını artırmak ve ülkedeki genel refah seviyesini yükseltmek için özellikle kadınların yükseköğrenime teşvik edilmesinin kritik bir öneme sahip olduğunu savunmuştur.

a) Ekonomistin bu iddiasının dayanağı nedir? Grafikteki verileri kullanarak açıklayınız.

b) Grafiğe göre, eğitim seviyesi ile işsizlik oranı arasında nasıl bir ilişki vardır? Bu ilişki erkekler ve kadınlar için benzerlik gösteriyor mu?

a) Dayanak: Tabloya göre kadınlarda okuryazar olmayanların işsizlik oranı %25.8 iken, üniversite mezunlarında bu oran %13.4'e düşmektedir. Eğitim seviyesi arttıkça işsizlik oranındaki düşüş kadınlarda çok daha belirgindir.

b) İlişki: Eğitim seviyesi arttıkça işsizlik oranı azalır (ters orantı). Bu ilişki her iki cinsiyet için de geçerlidir ancak kadınlarda oranlar genel olarak erkeklerden daha yüksektir

Örnek...11 :

Bir belediye, şehir içi ulaşımında bisiklet kullanımını teşvik etmek ve bisiklet yollarındaki güvenliği artırmak amacıyla yeni bir uygulama başlatmıştır. Hız sınırınının 20 km/sa. olduğu belirlenmiş pilot bir bisiklet yoluna, hız ihlallerini tespit etmek ve bisikletlileri kurallara uymaya teşvik etmek için "Hız Bilgilendirme Ekranları" (HBE) yerleştirilmiştir. Bu ekranlar, anlık hızını geçen bisikletlilerin hızını yansıtarak onları bilgilendirmeyi amaçlamaktadır.

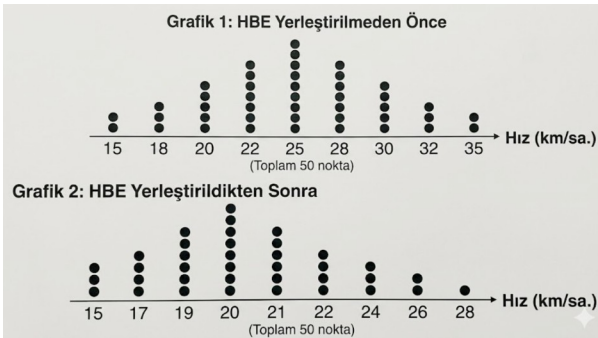
HBE'lerin yerleştirilmesinden önce 40 bisikletlinin ve HBE'lerin yerleştirilmesinden sonra 40 bisikletlinin hızları ölçülmüş ve aşağıdaki nokta grafiklerinde gösterilen veriler elde edilmiştir. Bisikletlilerin hızlarını düşürmelerinde HBE'lerin etkisinin olup olmadığı araştırılmıştır. Gözlemler, HBE'ler yerleştirildikten sonra bisikletlilerin hız sınırına daha çok dikkat ettiklerini göstermektedir.

Grafik 1: HBE Yerleştirilmeden Önce

- Yatay eksen (Hız - km/sa.): 15, 18, 20, 22, 25, 28, 30, 32, 35
- Nokta sayıları sırasıyla: 2, 4, 6, 8, 10, 8, 6, 4, 2 (Toplam 50 nokta)

Grafik 2: HBE Yerleştirildikten Sonra

- Yatay eksen (Hız - km/sa.): 15, 17, 19, 20, 21, 22, 24, 26, 28
- Nokta sayıları sırasıyla: 4, 6, 8, 10, 8, 6, 4, 3, 1 (Toplam 50 nokta)



Buna göre, HBE'lerin bisikletlilerin hızları üzerindeki etkisini istatistiksel yöntemleri kullanarak açıklayınız.

Yerleştirilmeden Önce: En çok tekrarlanan hız (Mod) 25 km/sa'dir ve 50 kişiden 34'ü hız sınırını (20 km/sa) aşmıştır.

Yerleştirildikten Sonra: En çok tekrarlanan hız (Mod) 20 km/sa'ye düşmüştür. Veriler sola doğru (düşük hızlara) kaymıştır.

Sonuç: HBE'ler bisikletlilerin hızlarını düşürmede ve yasal sınıra uymalarında etkili olmuştur.